

## 因子分析から主成分分析へ

勝 浦 正 樹

### はじめに

主成分分析は、アメリカの統計学者 Harold Hotelling によって提案された多変量解析の方法である。主成分分析が初めて示された論文 Hotelling [14] において、主成分分析は、因子分析の1つの手法として示されている。しかしながら、現在では、因子分析と主成分分析は区別して考える方が一般的である。<sup>(1)</sup> そこで、本論文では、因子分析の1つとして考案された主成分分析が、どのような歴史的変遷を経て、因子分析から独立していったのかを考察していこうと思う。

以下では、まずはじめに、両者の類似点、相違点を形式的な側面からまとめてみる。それらをふまえたうえで、主成分分析が、いつ、どのようにして因子分析から分離していったのかに関する1つの考え方を示す。そして、その背景となった主成分分析をめぐるいくつかの論争を類型化してみようと思う。

注(1) 現在の代表的な多変量解析のテキストのほとんどが、両者を別個に（章を分けて）扱っている。たとえば、Anderson [4], Karson [18], 竹内・柳井 [35], 田中・脇本 [36] など。

### 1. 因子分析と主成分分析の形式的同等性

因子分析における因子の抽出法には、いろいろな方法が考えられている。そのうち、主因子法といわれる方法のなかの特殊な方法が、主成分分析とはまっ

たく形式的には同じであることを簡単に示しておこう。

(1) 主因子法

いま、 $N$  個の観測値が、それぞれ  $p$  個の変数について測定されているとする ( $X_{ij}; i=1, \dots, N, j=1, \dots, p$ )。ここで、 $X_{ij}$  を平均と標準偏差で基準化した変数を  $z_{ij}$  とする。そこで、因子分析のモデルは、

$$z_{ij} = \beta_{1j}f_{i1} + \beta_{2j}f_{i2} + \dots + \beta_{mj}f_{im} + v_{ij} \quad (1)$$

と表される。 $f$  は共通因子 ( $m$  個)、 $\beta$  は因子負荷量、 $v$  は個別因子と呼ばれる。つまり、因子分析とは、与えられた観測値  $z_{ij}$  から、これらを推定する方法である。

推定のために、 $f, v$  には、次のような仮定が置かれる。

$$\left. \begin{aligned} E(f_j) &= 0, \text{ var } (f_j) = 1, \\ E(v_j) &= 0, \text{ var } (v_j) = s_j^2, \\ \text{Correl } (f, v) &= 0, \text{ Correl } (v_{ij}, v_{il}) = 0 \ (j \neq l), \\ \text{Correl } (f_{ik}, f_{il}) &= 0 \ (k \neq l) \end{aligned} \right\} \quad (2)$$

したがって、(1)、(2) より、

$$\text{var } (z_j) = \beta_{1j}^2 + \dots + \beta_{mj}^2 + s_j^2 = 1 \quad (3)$$

$$\text{Correl } (z_j, z_l) = r_{jl} = \beta_{1j}\beta_{1l} + \dots + \beta_{mj}\beta_{ml} \quad (4)$$

が得られる。行列表示すれば、

$$R = \frac{1}{N} Z'Z = B'B + S \quad (5)$$

$R$ : 相関行列 ( $p \times p$ )

$B$ : 因子負荷量行列 ( $m \times p$ )

$S$ :  $s_j$  を対象要素にもつ対角行列 ( $p \times p$ )

となる。よって、 $R - S = B'B$  となり、まず  $S$  を推定することになる。推定された  $S$  の対角要素により、 $1 - s_j^2 = \beta_{1j}^2 + \dots + \beta_{mj}^2$  が計算できるが、これを共通性という。何らかの方法<sup>[2]</sup> で共通性を推定し、 $R - S$  をつくれば、あとは、それを対角化するような、直交行列  $B$  を計算すればよい。つまり、 $R - S$

という行列に関する固有値問題を解けばよいことになる（ $m$  個の固有値が正であるとする）。このようにして、因子負荷量  $\mathbf{B}$  を決定する方法が、主因子法である。

## (2) 主成分分析

主因子法の大きな問題点は、共通性の推定と因子数  $m$  の決定に恣意性が入り易いことである。そこで、共通性を推定せず、 $\mathbf{R}-\mathbf{S}$  ではなく、相関行列  $\mathbf{R}$  を直接対角化しようとするのが、主成分分析である。つまり、共通性の要素をすべて 1 とし、 $\mathbf{R}=\mathbf{B}'\mathbf{B}$  という固有値問題を解くことになる。よって因子は、 $\mathbf{R}$  の次数  $p$  だけ抽出されることになる。<sup>(3)</sup> つまり、主成分分析は、相関行列の対角要素に共通性の推定値を置かず（共通性の推定における恣意性は排除される）、共通性を 1 のままにして因子を抽出する主因子法ということができる。

このことから、主成分分析のモデルは、形式的には、

$$z_{ij}=\beta_{1j}f_{i1}+\beta_{2j}f_{i2}+\cdots+\beta_{pj}f_{ip} \quad (6)$$

と書くことができる。

## (3) 主因子法と主成分分析の違い

以上のように、形式的には、主成分分析を因子分析（主因子法）の枠の中でとらえることができるが、両者の違いを簡単にまとめてみよう。

(1) と (6) を比べれば明らかなように、主成分分析では、共通因子と個別因子を区別せず、個別の因子をモデルに組み込んでいない。したがって、主成分分析では、必然的に  $p$  個の因子（主成分）が抽出されてしまう。このことは、少数の因子でデータを説明しようとする因子分析の発想とは異なることになる。

ところで、通常の変量解析の文献においては、<sup>(4)</sup> 主成分分析における主成分の抽出は、(6) のようなモデルから出発しない。そこにおいては、むしろ、

$$f_{ij}=\beta_{j1}z_{i1}+\beta_{j2}z_{i2}+\cdots+\beta_{jp}z_{ip} \quad (7)$$

といった、第  $j$  主成分  $f_j$  を考えて、その  $f$  の分散が最大のものから、第 1 主成分、第 2 主成分、…、第  $p$  主成分と呼び、抽出していく。そのような基準

から出発しても、結局は、 $z$  の相関行列  $R$  を対角化するという固有値問題に帰着するのである。

ここで、(1) や (6) とまったく異なっているのは、(7) は、 $p$  個の変数を因子負荷量（固有ベクトル）をウェイトにして総合化し、新たな変量  $f$  を合成するという発想である。つまり、 $z$  を  $f$  に分解するのが因子分析で、 $z$  から  $f$  を合成するのが主成分分析であるといえよう。<sup>[5]</sup> つまり、形式的には、両者は同一の手法としてまとめることができるが、それらの方法の基本的な目的・方向はまったく異なっているのである。

注(2) 共通性の推定方法には、 $R$  の第  $j$  要素のうち絶対値が最大のものを用いる方法、 $z_j$  を他の  $p-1$  個の  $z$  で回帰したときの決定係数を用いる方法、反復的に計算していく方法、共通性を直接推定しない方法などが考えられている。岩田 [17] 271—273 ページ、Comrey [12] などを参照のこと。

(3)  $R$  は、正値定符号行列である。主成分分析では、相関行列  $R$  ではなく、 $X_{ij}$  の分散・共分散行列を  $B$  によって、対角化するのが通常である。しかし、測定単位の異なる  $X_{ij}$  の場合、基準化することが多く、その場合、基準化した変数の分散・共分散行列は、もとの変数の相関行列に一致するので、 $R$  を対角化することになる。

(4) 注(1)であげた文献など。

(5) Lawley and Maxwell [23] 邦訳 3—4 ページ、田中・脇本 [36] 180 ページなど多くの文献で指摘されている違いである。

## 2. 因子分析と主成分分析の歴史的変遷

前節で述べたように、因子分析と主成分分析は、その目的が大きく異なっており、現在では、区別して用いられることが多い。しかし、主成分分析も、Hotelling [14] において初めて提案された時点では、因子分析の1つの手法として示されていた。それが、どのような経過をたどって、現在のように両者が区別されるようになったのかを、この節では歴史的に見ていくことにする。

### (1) 因子分析の起こり

因子分析という方法が初めて示されたのは、Spearman [33] においてであ

とされている。<sup>[6]</sup> Spearman がそうであったように、以後、因子分析は、心理学への応用を中心に発達してきた。そして、1920年代頃までに、Spearman, Burt, Thomson, Garnett らによって、心理学における統計手法として確立するに至った。<sup>[7]</sup> 因子分析は、人間の能力を測定し、そのためにテスト得点などを分析するいわゆる計量心理学（Psychometrics）の発展と深い関わりをもっている。Spearman [33] における一般因子・個別因子説（2因子説）から、Burt の総和法、Kelly の主軸法、そして、Thurstone の多因子説などへと到る。つまり、その初期の発展の大部分は、心理学者の手に委ねられていたのである。

## （2）主成分分析の起こりと展開

主成分分析は、1933年の Hotelling [14] において、因子分析の1つの方法として提案された方法である。Hotelling は、因子分析のモデル（1）をより一般的に、（6）のように示し、 $z$  の分散への寄与の大きい順に、 $\beta$  や  $f$  を決定していく方法を考えた。Hotelling は、その方法を method of principal component と呼び、 $f$  を因子（factor）ではなく、成分（component）と名付けた。そして、より因子分析的に、 $z$  の分散への貢献が小さい成分は無視するべきであると主張している。また、現在用いられている主成分分析と大きく異なっているのは、 $\beta$  は、 $R$  の固有ベクトルそのものではなく、固有値の平方根で基準化することにより、因子負荷量に一致するようにしており、主成分の分散が1になるようにしていることである。<sup>[8]</sup>

さて、数理統計学者である Hotelling が、心理学の範囲にあった因子分析を、なぜ研究対象としたのかというのは興味深い問題である。<sup>[9]</sup> しかし、ここではその問題には立ち入らずに、因子分析に統計学者の議論が投げ掛けられたことが、その後の因子分析等の研究にどのような影響をもたらしたのかを中心にみていこう。

まず、Hotelling の主成分分析に批判を加えたのは、アメリカの計量心理学者 Thurstone であった（1935年）。<sup>[10]</sup> Thurstone の批判の根拠は、主成分分

析では、 $p$  個の成分が抽出されてしまい、データに潜む単純構造を見付け出すのには役立たず、因子分析の目的に合わないというものである。<sup>[41]</sup> この Hotelling の主成分分析の考えと Thurstone の批判は、その後1930年代後半から40年代<sup>[42]</sup>における、主成分分析と因子分析をめぐるかなり活発な議論の発端となったようである。続いて、1936年には、統計学者 Girshick が [13] において、Thurstone の Hotelling への批判を和らげ、Hotelling を擁護するような論文を出している。<sup>[43]</sup> その後、主成分分析自体は、心理学者の立場からは批判的に見られていたが、他の分野で応用されていった。たとえば、統計学では、Kendall [19] (1939)、経済学では、Tintner [39] (1946)、Stone [34] (1947) などである。これらの応用においては、主成分分析は、因子分析的なデータの変動要因の解明とともに、主成分得点という形で変数を合成するための手法としても利用されている。<sup>[44]</sup>

### (3) 因子分析と主成分分析の対立と分離

1940年代には、因子分析自体においても、様々な議論が展開された。特に、Hotelling 以来、心理学者だけではなく、統計学者たちによって、その手法に改良が加えられたり、因子分析の目的、問題点といったものが論じられるようになったのである。たとえば、1940年代の始めには、Lawley によって、最尤推定法が考えられているし、<sup>[45]</sup> Bartlett も、[6][7] などにおいて、因子分析の特徴・目的などを統計学的な立場から議論したり、標本理論を取り入れようとしていたりしている。

心理学独自のテリトリーにあった因子分析を、統計学者が扱うようになって、因子分析も数学的に高度になってきたし、発展したことは事実である。しかし一方で、心理学者と統計学者の対立も表面化するようになってきたのである。それが最も顕著に現れているのが、主成分分析と因子分析の対立なのである。アメリカでは、先に述べたように Thurstone が Hotelling の主成分分析を批判し、因子分析こそが重要であるとしているし、イギリスにおいても、同様の批判があった。

ところで、先に因子分析は、Spearmanによって最初に導入されたと述べたが、イギリスの心理学者 Burt は [10] において、因子分析を最初に考案したのは、Pearson であると解釈している。Burt は、その根拠として 1901 年の Pearson [29] という論文を指摘している。<sup>66</sup> Pearson の論文は、散布図に直線をあてはめるときに、通常の最小 2 乗法（各点からあてはめる直線に、 $y$  軸に平行に下ろした線分の長さの平方和の最小化）ではなく、あてはめる直線に下ろした垂直の長さの平方和が最小になるようにする方法を示したものである。この方法は、主成分分析または主軸法を幾何学的に解釈したものであり、それらの手法とまったく一致する。<sup>67</sup> しかし、Pearson は、あくまでも直線のあてはめを目的としていたのであって、変数間の共通因子（成分）を求めようとしたのではない。やはり、因子分析の origin は、Spearman に求めるべきであろうと思われる。

このように、主成分分析をふくめた因子分析については、誰がはじめに提案したかという議論まで出てきて、これらの手法をめぐる対立はかなり顕在化し、また混沌としてきた。単純化すれば、心理学者は主成分分析を批判の対象とし、統計学者は主成分分析を擁護するといった構造である。主成分分析は一意的に  $k$  個の成分が得られるという客観的なものであるということが、統計学者が因子分析よりも主成分分析を好んだとも考えられている。<sup>68</sup> そして、しだいに主成分分析は、因子分析から独立していったものと思われる。つまり、因子分析を心理学者だけでなく、統計学者が研究の対象とするとともに、因子分析は心理学者・統計学者に共通の研究対象となったが、因子分析の一部であった主成分分析は、むしろ、心理学から離れ、応用範囲も他の分野へと進出していった。また、主成分分析という独自の手法としての研究自体が、統計学者の手に委ねられていった。

このことが、明確に見られるのが、1950 年の Kendall and Smith [21] における Kendall の論文 “Factor Analysis as a Statistical Technique” であろう。ここにおいて、主成分分析はデータに潜む共通の変動要因の抽出を目

的とするといった因子分析的発想より、変数の総合化というところに重点が置かれている。

「成分分析によって扱われる一般的な問題は、次の通りである。……もとの変数を線形的に結合させた新たな変数を、しかももとの変数より少ない数で、見つけることができるか？……成分分析の基本的な問題は、変数の数を節約することであり、新たな変数が、“real”なものとして識別できるかが、望まれる目標であると言えるだろう。」<sup>[19]</sup>

また、この論文や Stone [34], Bartlett [5] などの論文に含まれる Discussion では、統計学者・心理学者の両方から、様々な議論が出されており、両者の対立は明確になっている。<sup>[20]</sup>

さらに、1950年代にも、様々な議論があったようであるが、<sup>[21]</sup> 主成分分析が因子分析から、ほぼ完全に独立したのは、Kendall [20] (1957) においてであると思われる。この著は、初めての多変量解析の総合的なテキストであるが、主成分分析と因子分析は別個に扱われている。もちろん、主成分分析の目的と因子分析の目的は、明確に区別されている。<sup>[22]</sup> つまり、主成分分析は、Hotelling が示したような因子分析的モデル（6）ではなく、（7）というモデルを前提に置き、<sup>[23]</sup> 総合指標の作成が主目的であるとしており、議論を展開している。特に、各変数を総合化した主成分が独立であるという性質を利用して、回帰分析において多重共線性を回避するために主成分回帰の方法を提案している。<sup>[24]</sup> 主成分回帰は、因子分析的な側面がまったく見られない、主成分分析に独特な利用方法であり、このあたりからも、主成分分析が因子分析から独立している様子をみることができよう。また、Hotelling [14] では、（7）の  $\beta$  は因子負荷量となっているが、Kendall においては、固有値で基準化せず、 $R$  の固有ベクトルをそのまま用いており、主成分の分散は1にされていないということも、主成分分析が因子分析と区別される上で興味深い点である。

さらにこれとは別の流れとして、1958年に Anderson [1] が出版されている。この名著でも、因子分析と主成分分析は、はっきりと区別され、特に、数



学的な明確さをもって示されているのが大きな特徴である。主成分の導出の仕方、総合指標的な立場によって、主成分の分散を最大化するという最もポピュラーな方法が採用されている。主成分分析が、数学的な特性からみても統計学者にとって興味深いということが現れており、これ以後、Anderson [3] などのように、推測統計的に主成分分析を考察する方向へと向かっていくのである。

以上のように、主成分分析が純粹に心理学的な手法ではなく、心理学者から統計学者の手で扱われるようになるとともに、その分析目的も因子分析とは区別されていき、主成分分析は因子分析から独立していったといえよう。そして、そこで最も大きな役割を果たしているのは、Anderson が [1] を著すのに Kendall [20] を参考にしたのかどうかは別にしても、やはり、Kendall であるといえよう。Kendall の貢献によって、現在のような多変量解析の 1 手法としての主成分分析が確立したと考えられる。<sup>24)</sup>

注(6) 後に見るように、Burt などは、Galton や Pearson に、因子分析の発端を見出している。しかし、現在では、一般的には、Spearman の 1904 年の論文が因子分析の始まりだと言われている。

(7) 田中 [37] 175—176 ページ、中邑 [26] i—ii ページ。

(8) Hotelling [14] pp. 417-426. 現在では、主成分の分散を 1 にせず、その分散（固有値に一致する）により、その主成分の寄与を表わすのが通常である。

(9) Hotelling は、数理統計学者として、主成分分析のほかに、Hotelling の  $T^2$  や正準相関分析などで有名である。その他にも、アメリカの統計教育や数理経済学（marginal cost pricing の研究など）への貢献が大きく、アメリカ統計学会においても多大な影響力をもっていた。また、Hotelling に最も大きな影響を与えたのは、R. A. Fisher だと言われている。Hotelling の経歴・業績などについては、Anderson [2], Levene [24], Madow [25], Olkin et al. [27], Samuelson [31], Smith [32] などを参照のこと。

(10) Thurstone の著 *Vectors of Mind* で展開されており、Girshick [13] pp. 519-520 でまとめられている。

(11) Thurstone [38] pp. 503-510 など。

(12) 1936 年の *Psychometrika*, 1947 年の *British Journal of Psychology* の Statistical Section の発刊にみられるように、計量心理学が、著しい発展を見せた時期

でもある。本論文の巻末の参考文献以外にも、これらの雑誌に、多くの計量心理学者の論文が掲載されており、因子分析の発展の経過・論争の過程をよりくわしくみることができる。

- (13) Girshick [13] p. 520. この論文は、Hotelling の方法を数理的に分かりやすく示し、Thurstone の批判が適切でないことを述べている。さらに、因子分析（主成分分析）において、最尤法の概念を導入しようとしている。
- (14) たとえば、Tintner [39] pp. 482-489 では、いくつかの生産指数や物価指数を用いて、それらを総合し、新たな生産指数や物価指数を作成しようと試みている。
- (15) 因子分析における最尤推定は、Lawley [22] で提案され、Lawley and Maxwell [23] (1963) でまとめられている。
- (16) Burt は、この論文以外にもこのことを頻繁に主張している。Kendall and Smith [21] に含まれる Discussion (pp. 87-88) でも、因子分析における crucial article は、Spearman [33] (1904) や Hotelling [14] (1933) ではなく、Pearson [29] (1901) であり、Pearsonこそが、提案者だとしている。その他に、Stone [34] p. 36, Bartlett [5] p. 192 に含まれる Burt の Discussion などにみられる。
- (17) たとえば、Kendall [20] 邦訳 8—11 ページや奥野他 [28] 174—175 ページなどで証明されている。垂線の平方和の最小化という基準は、主成分導出の 1 つの基準であり、分散の寄与の大きい主成分から抽出していく基準、各変数と主成分との重相関係数の平方和の最大化といった基準と同値である。
- (18) Kendall and Smith [21] p. 60.
- (19) Kendall and Smith [21] p. 61.
- (20) Kendall and Smith [21] pp. 87-88 の Burt の意見などで、心理学者の立場からみた因子分析というのは、“confirmatory” であるとしている。また、統計学者に対して、因子分析における有意性検定の必要ということを要求している。いずれにせよ、統計学者に対しては、いくぶん批判的である。
- (21) たとえば、1953 年に行われた Uppsala Symposium on Psychological Factor Analysis など (Bartlett [8], Peel [30])。
- (22) Kendall [20] 邦訳 29 ページ。
- (23) Kendall [20] 邦訳 5 ページ, 10 ページ。
- (24) Kendall [20] 邦訳 56—61 ページ。
- (25) Anderson [1] の reference では、Kendall [20] は、挙げられていない（ただし、Kendall and Smith [21] は、挙げられている）。
- (26) 総合指標的な立場を主張している論文には、主成分分析の応用として、Tintner [39] や Stone [34] があるが、それらに先がけて、Kendall が [19] で、総合指標的な利用を行っている。これからも、やはり、Kendall に最大の貢献があるほしい。

### 3. 主成分分析をめぐる対立の構造

前節で述べたような経過をたどって、主成分分析は因子分析から独立してきたが、その背景として、いくつかの学派の対立が存在していた。以下では、それらの対立・論争・批判の構造を単純化し、類型化しておこう。

まず、主成分分析・因子分析をめぐる対立の中心的な存在であった心理学者・統計学者をあげると表1のようになる。<sup>20)</sup> 当時の因子分析の研究においては、イギリスの心理学者は Burt に、アメリカの心理学者は Thurstone に従うとされている。<sup>21)</sup> したがって、彼らの意見はそのまま、それぞれの国での心理学の動向を反映していると考えてもよいだろう。表1の4つのブロックは、それぞれの主張をもって、他のブロックと対立していたのである。そこで、彼らの意見を中心にして、論争を見ていくことにしよう。

#### (1) 心理学者対統計学者

前節でもみたように、主成分分析・因子分析をめぐる心理学者と統計学者の対立は、かなり明確であった。特に、アメリカにおける Thurstone の主成分分析批判、Hotelling ら統計学者への批判はかなり強いものであった。

「因子分析の理論に貢献した数理統計学者は、主軸法の解とその特性の研究に限定してきた。これは、主軸法の解が数学的な取り扱いやすさをもっているので当然のことである。数学者たちが、単純構造の概念とその意味を展開するような問題にも目を向けることが望まれる。」<sup>22)</sup>

主成分分析（主軸法）は、心理学で用いられる因子分析のように、因子の数などに心理学的なモデルを与えずに、一意的に  $p$  個の主成分を抽出し、それ

表1 主成分分析・因子分析に関する論争における主な心理学者・統計学者

	イギリス	アメリカ
心理学	Burt	Spearman, Thurstone
統計学	Kendall, Bartlett	Hotelling, Anderson

が、因子分析本来の目的である少数個の因子をみつけ、データを説明することに反するという批判である。つまり、主成分分析はいくら客観性を重んじているとしても、心理学としては受け入れにくい手法ということであろう。この批判は、確かに妥当性をもっているが、Hotelling 自身、因子分析的発想で、 $p$  より少ないできるだけ少数の主成分を得ることが基本であるとしている。<sup>80)</sup> いずれにせよ、Thurstone は、主成分分析—統計学者、因子分析—心理学者という図式を描いていたように思われる。そして、主成分分析に因子分析的でない側面があることを批判の対象とし、Hotelling などの統計学者への批判につながったといえよう。<sup>81)</sup>

このような対立は、イギリスにおいてもみることができる。Kendall は、心理学者は因子分析を心理学の手法としてのみ利用可能としているが、そうではなく、より一般的な統計的手法として利用すべきであると述べている。また、因子分析がそれまで統計学者に受け入れられにくかったのは、心理学的専門用語が多過ぎて分かりにくく、客観性に欠けるものとして信用しにくかったとも述べている。<sup>82)</sup>

当時の様々な因子分析や主成分分析に関する文献では、「心理学者は……」、「統計学者は……」といった表現がかなり見られ、両者の対立はかなり強かったようである。しかし、これらの論争・対立を通じて、因子分析が、心理学固有の方法から、より広い応用分野へと広がっていったことは事実であり、その対立の接点に主成分分析があったという解釈も可能であろう。さらに、論争を通じて、因子分析の目的、主成分分析の目的がそれぞれ明確化され、両者は分離していった。そして、Bartlett が Uppsala Symposium on Psychological Factor Analysis で述べているように、心理学者と統計学者はそれぞれの専門を生かしながら、互いに協力していくほうがよいという意見も出されるようになったのである。<sup>83)</sup>

いずれにせよ、主成分分析・因子分析の客観性の問題は大きな論点となっていたのである。このことは、両者の発展に大きな影響を与えている。たとえば、

主成分分析は  $p$  個の主成分が得られてしまうという批判は、重要な主成分だけを検定によって客観的に選択するという Anderson の検定<sup>84</sup> などが生まれるに至った背景になったと言えるだろう。因子分析においても、より客観的なモデル選択や推定方法が開発されるようになった。<sup>85</sup>

(2) イギリス対アメリカ

このように心理学者と統計学者の論争（表1のたて方向の対立）は、論点がある程度はっきりしていて、それなりの成果をみることができるが、対立のもう1つの構造としては、イギリスとアメリカの対立をみることができる。とりわけ、これは、イギリスの心理学者から出されたものであった。

ここにおける論点は、主成分分析や因子分析の目的・方法論上の問題といったものではない。当時のイギリス計量心理学の権威であった Burt がさかんに問題にしているのは、誰が初めに因子分析を開発したかであった。前節でも述べたように、Burt は、因子分析も主成分分析も含めて Pearson にすべて貢献を帰している。因子分析—Spearman、主成分分析—Hotelling という構図を認めていない。確かに、Pearson は主成分分析の原形を出しているが、それは、回帰分析の1つとしてであり、因子分析・主成分分析とは別個に考えるべきであろう。これは、Burt ら心理学者のイギリス至上主義、あるいは、権威主義の現れであったとしか言い様がない。実際、Burt は、Hotelling の方法には批判的であったにもかかわらず、イギリスの統計学者 Kendall が、Pearson の方法は主成分分析と一致することを明確に証明したことに対しては、かなりの賞賛を与えている。<sup>86</sup>

このように、Burt は、心理学であれ、統計学であれ、アメリカのものを排除したかったようである。<sup>87</sup> 権威者 Burt のこの態度は、イギリスの心理学全体に強い影響を及ぼしたであろうことは、想像に難くない。<sup>88</sup> 因子分析の研究者であった Burt が、主成分分析を排除したいという理由は、Thurstone のような理由をもってしてではなかったのであり、こうした態度は、心理学と統計学という学問上の対立を混乱に陥らせ、論点を不明確にするものであったと

いえる。Thurstone の Hotelling 批判以来、主成分分析と因子分析の分離に約20年という長い時間を要した1つの理由をここにみることができよう。

統計学に関して言えば、この論争では、とりたててイギリスとアメリカの対立は見られない。<sup>38)</sup> 統計学者においては、Hotelling の主成分分析を、その客観性から、擁護する立場が共通のものであったようである。

注(27) 表1にあげた人物のほかにも、様々な学者がこの問題を論じている。たとえば、

Girshick, Harman, Holzinger, Lawley, Maxwell, Peel, Thomson 等である。

(28) 中邑 [26] ii ページ。

(29) Thurstone [38] p. 510. ここで用いられている主軸法と主成分分析とは、ほぼ同義である。

(30) Hotelling [14] pp. 417-422.

(31) こうした批判に対して、Hotelling は、特に反論等をしていないようである。

Hotelling は、主成分分析に関する2つの論文 (Hotelling [14] [15]) 以後、主成分分析についての論文を出していない。後に多変量解析の動向について述べた[16]においても、主成分分析については、あまり触れられていない。

(32) Kendall and Smith [21] p. 60.

(33) Bartlett [8] pp. 23-24.

(34) Anderson [3].

(35) たとえば、Bartlett [6] [7] や Lawley and Maxwell [23] など。

(36) Kendall and Smith [21] pp. 87-88.

(37) Burt [10] をはじめ、Burt の論文の reference には、Thurstone などアメリカの学者の名があまり挙げられていないことにも、この傾向をみることができる。

(38) Kendall and Smith [21] p. 88 の Burt の Smith に対する Discussion において、「この国 (イギリス) では、主軸法を提案したことに対する名誉を Hotelling ではなく、Pearson に与えるべきである。Hotelling は、(Pearson が提案してから) 22年後 (32年の誤り?) になるまで、主軸法について示していない」と述べている。Spearman についても同様に、彼の因子分析の提案者としての貢献を認めていない。

(39) Kendall が、主成分分析における主成分の導出に、Pearson 流の垂線の平方和の最小化をよく用いていることに、やや、イギリス重視が見られないわけではないが、特に、Hotelling などを批判していることはない。

## おわりに

以上のように、心理学における因子分析の1つの方法として、Hotelling によって登場した主成分分析は、心理学者と統計学者の論争を経て、因子分析から独立していった。それは、1950年代にみることができ、イギリスの統計学者 Kendall が、それに大きな役割を果たしているという1つの考えをここに提示した。特に、独立の過程で、分析の目的が、Hotelling の因子分析的な目的とは違い、主成分という総合的指標の作成に重点が置かれるようになったことは、大きな転換点となったのである。

このような歴史的展開をみることによって、主成分分析という方法の利用の仕方、結果の解釈に注意しなければならないことが感じられる。現在の応用研究では、主成分分析を因子分析的に用いて、変動要因の抽出という点に重点が置かれる場合も多い。もちろんそれが誤りであるとはいえないが、主成分分析の主要目的は、あくまでも、変数の変動をうまく説明するような総合的指標である主成分を計算することにあり、変動要因としてあまりにもそれを信頼・過信することは危険である。変動要因の抽出ということだけからすれば、因子分析の方が目的にあっていいる。因子分析を用いるのか、主成分分析を用いるのかを、分析の目的や状況に応じて適切に使い分けることが、重要な点である。Bartlett は、Uppsala Symposium on Psychological Factor Analysis で次のように述べている。<sup>(40)</sup>

「同じデータでも、目的が異なれば、異なった方法で、しかもうまく特定化されうるだろう。」

注<sup>(40)</sup> Bartlett [8] p. 24.

## 参考文献

- [1] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, 1st ed., John Wiley & Sons, 1958, 374 pp.
- [2] Anderson, T. W., "Harold Hotelling's Research in Statistics," *The Ameri-*

- can Statistician*, vol. 14, June, 1960, pp. 17-21.
- [3] Anderson, T. W., "Asymptotic Theory for Principal Component Analysis," *Annals of Mathematical Statistics*, vol. 34, 1963, pp. 122-148.
  - [4] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, 2nd ed., John Wiley & Sons, 1984, 657 pp.
  - [5] Bartlett, M. S., "Multivariate Analysis," (with Discussion) *Supplement to the Journal of the Royal Statistical Society*, vol. 9, 1947, pp. 176-197.
  - [6] Bartlett, M. S., "Internal and External Factor Analysis," *British Journal of Psychology (Statistical Section)*, vol. 1, part II, 1948, pp. 73-81.
  - [7] Bartlett, M. S., "Tests of Significance in Factor Analysis," *British Journal of Psychology (Statistical Section)*, vol. 3, part II, 1950, pp. 77-85.
  - [8] Bartlett, M. S., "Factor Analysis in Psychology as a Statistician Sees It," in *Uppsala Symposium on Psychological Factor Analysis*, Nordisk Psykologi's Monograph Series No. 3, 1953, pp. 23-34.
  - [9] Burt, C., "A Comparisom of Factor Analysis and Analysis of Variance," *British Journal of Psychology (Statistical Section)*, vol. 1, part I, 1947, pp. 3-26.
  - [10] Burt, C., "Alternative Methods of Factor Analysis and Their Relations to Pearson's Method of 'Principal Axes'," *British Journal of Psychology (Statistical Section)*, vol. 2, 1949, pp. 98-121.
  - [11] Burt, C., "Group Factor Analysis," *British Journal of Psychology (Statistical Section)*, vol. 3, part II, 1950, pp. 40-75.
  - [12] Comrey, A. L., *A First Course in Factor Analysis*, Academic Press, 1973 (芝祐順訳『因子分析入門』サイエンス社 1979年 239ページ)。
  - [13] Girshick, M. A., "Principal Components," *Journal of the American Statistical Association*, vol. 31, 1936, pp. 519-528.
  - [14] Hotelling, H., "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology*, vol. 24, 1933, pp. 417-441 and 498-520.
  - [15] Hotelling, H., "Simplified Calculation of Principal Components," *Psychometrika*, 1936, vol. 1, pp. 27-35.
  - [16] Hotelling, H., "The Relations of the Newer Multivariate Statistical Methods to Factor Analysis," *The British Journal of Statistical Psychology*, vol. 10, part II, 1957, pp. 69-79.
  - [17] 岩田暁一『計量経済学』有斐閣 1982年, 364ページ。
  - [18] Karson, M. J., *Multivariate Statistical Methods*, Iowa State Univ., 1982,



307 pp.

- [19] Kendall, M. G., "The Geographical Distribution of Crop Productivity in England," (with Discussion) *Journal of the Royal Statistical Society*, vol. 102, 1939, pp. 21-62.
- [20] Kendall, M. G., *A Course in Multivariate Analysis*, 1957, Charles Griffin, 1957, 185 pp (浦昭二・竹並輝之訳『多変量解析の基礎』サイエンス社 1972年, 160ページ).
- [21] Kendall, M. G. and Smith, B. B., "Factor Analysis," (with Discussion), *Journal of the Royal Statistical Society* (Series B), vol. 12, 1950, pp. 60-94.
- [22] Lawley, D. N., "The Estimation of Factor Loadings by the Method of Maximum Likelihood," *Proceeding of the Royal Society of Edinburgh*, vol. 40, 1940, pp. 64-82.
- [23] Lawley, D. N., and Maxwell, A. E., *Factor Analysis as a Statistical Method*, Butterworth, 1963 (丘本正監訳『因子分析法』日科技連 1970年, 163ページ).
- [24] Levene, H., "Harold Hotelling 1895-1973," *The American Statistician*, vol. 28, May, 1974, pp. 71-73.
- [25] Madow, W. G., "Harold Hotelling as a Teacher," *The American Statistician*, vol. 14, June, 1960, pp. 15-17.
- [26] 中邑幾太『因子分析学』日本文化科学社 1987年, 182ページ。
- [27] Olkin, I. et al. ed., *Contributions to Probability and Statistics*, Stanford Univ. Press, 1960, 517 pp.
- [28] 奥野忠一他『多変量解析法』〈改訂版〉日科技連 1981年, 430ページ。
- [29] Pearson, K., "On Lines and Planes Closest Fit to Systems of Points in Space," *Philosophical Magazine*, vol. 2, 1901, pp. 559-572.
- [30] Peel, E. A., "Factor Analysis as a Psychological Technique," in *Uppsala Symposium on Psychological Factor Analysis*, Nordisk Psykologi's Monograph Series No. 3, 1953, pp. 7-22.
- [31] Samuelson, P. A., "Harold Hotelling as Mathematical Economist," *The American Statistician*, vol. 14, June, 1960, pp. 21-25.
- [32] Smith, W. J., "Harold Hotelling," in *Encyclopedia of Statistical Sciences*, John Wiley & Sons, 1983, pp. 668-669.
- [33] Spearman, C., "General Intelligence Objectively Determined and Measured," *American Journal of Psychology*, vol. 15, 1904, pp. 201-293.
- [34] Stone, R., "On the Interdependence of Blocks of Transactions," (with Discussion) *Supplement to the Journal of the Royal Statistical Society*, vol.

9, 1947, pp. 1-45.

[35] 竹内啓・柳井晴夫『多変量解析の基礎』東洋経済新報社 1972年, 326ページ。

[36] 田中豊・脇本和昌『多変量解析法』現代数学社 1983年, 296ページ。

[37] 田中良久編『講座 心理学 2 計量心理学』東京大学出版会 1969年, 279ページ。

[38] Thurstone, L. L., *Multiple Factor Analysis*, The Univ. of Chicago Press, 1947, 535 pp.

[39] Tintner, G., "Some Applications of Multivariate Analysis into Economic Data," *Journal of the American Statistical Association*, vol. 41, 1946, pp. 472-500.

1989.3.30 提出

(博士後期課程第4年度生)